

MAP Based Iterative Channel Estimation for OFDM Systems: Approach, Convergence, and Performance Bound

Yinsheng Liu, Geoffrey Ye Li, Hongjie Hu, and Zhenhui Tan

Abstract—*Iterative channel estimation (ICE) usually exploits soft information of unknown data symbols as references to improve estimation performance. This paper investigates ICE for orthogonal frequency division multiplexing (OFDM) over wireless channels. The optimum ICE is derived in terms of maximum a posteriori (MAP) criterion, which can be solved using fixed-point iteration (FPI). Furthermore, the derived MAP ICE is closely related to the well-known expectation-maximization (EM) estimation. We also demonstrate that the MAP ICE converges within only one step when the signal-to-noise ratio (SNR) is large through analysis and simulation results.*

Index Terms—OFDM, iterative channel estimation, EM algorithm, performance bound.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) can provide high data rate transmission over wireless channels due to its robustness against frequency selectivity [1]. In an OFDM system, the frequency selective channel is converted into a group of flat fading channels, which can be easily compensated using a simple one-tap equalizer. As a result, the design of OFDM receiver is greatly simplified, and thus OFDM is widely used in many wireless systems [2].

With knowledge of frequency response of wireless channels at the receiver, coherent detection can be used in OFDM systems, which has a much better performance than differential detection that requires no channel knowledge. Channel parameters need to be estimated for coherent detection. Traditional approaches adopt one-round manner for channel estimation [3], [4], and, thus the estimation performance depends on the pattern of pilots [5].

Following the success of turbo code [6], the *turbo principle* has been widely applied in receiver design [7], and has been

first used in OFDM receiver in [8]. The turbo receiver enables channel estimation to be performed iteratively using the soft information of data symbols. With the soft information fed back by the detector, the unknown data symbols are replaced by the soft symbols which are the averages of possible constellation points. The regenerated soft symbols and the pilot symbols are used together to improve channel estimation for the next iteration. By taking into account the inaccuracy of regenerated soft symbols, an optimized iterative channel estimation has been developed in [9].

Since the regenerated soft symbols are not reliable when the *signal-to-noise ratio* (SNR) is low, different threshold based approaches have been proposed in [10]–[12] to improve the estimation performance in this case. The basic idea for the threshold approach is that a regenerated soft symbol is used for channel estimation only if its reliability is above a threshold. For a system with high order *quadrature-amplitude-modulation* (QAM), such as 64-QAM, the constellation contains many points, and thus regenerating soft symbols needs heavy computation. To reduce the complexity, a sub-constellation, such as 4-QAM, is used in [13], where the data symbol is regenerated using only several points with the highest probabilities. It is also possible to feed back the hard-decision symbols rather than the soft information [14].

In this paper, we propose an optimum *iterative channel estimation* (ICE) in terms of *maximum a posteriori* (MAP) criterion, which can be solved iteratively using *fixed-point iteration* (FPI) [15]. Instead of using arithmetic average [8]–[10], the derived optimum estimator suggests using a harmonic average over possible constellation points to regenerate soft symbols. We find a close connection between the proposed MAP ICE and the well-known *expectation-maximization* (EM) estimation [16], [17]. EM estimation includes two steps: an *expectation-step* (E-step) and a *maximization-step* (M-step). The EM estimation can find the *maximum-likelihood* (ML) estimation of unknown parameters through iterations between E-step and M-step when there is missing data. The EM estimation can be extended to MAP estimation by modifying the M-step with a given a priori information. Moreover, we demonstrate the monotonic convergence of the proposed MAP ICE, and show the proposed approach can converge within one step when the SNR is large, as confirmed by our analysis and the simulation.

The rest of this paper is organized as follows. The system model is presented in Section II. Optimum ICE is derived

Manuscript received June 8, 2013; revised August 21 and October 20, 2013; accepted October 23, 2013. The associate editor coordinating the review of this paper and approving it for publication was Z. Wang.

This work was supported by the Fundamental Research Funds for the Central Universities under Grant number 2013YJS022, and was supported by the National Science Foundation of China [No.61032002, No. 61071075 and No. 61001071].

Y. Liu and Z. Tan are with the Institute of Broadband Wireless Mobile Communications & State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China (e-mail: 09111035@bjtu.edu.cn).

G. Y. Li is with ITP Lab, School of ECE, Georgia Institute of Technology, Atlanta 30313, Georgia, USA (e-mail: liye@ece.gatech.edu).

H. Hu is with Huawei Technologies Co., Ltd., Shanghai, China (e-mail: hulongjie@huawei.com).

Digital Object Identifier 10.1109/TWC.2013.112513.131038

in Section III. Convergence behavior is discussed in Section IV. In Section V, the performance bound in terms of Bayesian *Cramer-Rao Bound* (CRB) [18] is analyzed. Simulation results are shown in Section VI. Finally, conclusions are drawn in Section VII.

II. SYSTEM MODEL

In this section, we will describe the channel model and the OFDM system with soft-output detector.

A. Channel Model

The complex baseband representation of *channel impulse response* (CIR) of a wireless channel can be described by

$$h(\tau) = \sum_l \alpha_l g(\tau - \tau_l), \quad (1)$$

where τ_l and α_l are the delay of the l -th path and its complex amplitude, and $g(\tau)$ is the impulse response of the equivalent system of the transmit and receive filter. Then, the *channel frequency response* (CFR) can be expressed as

$$H(f) = G(f) \sum_l \alpha_l e^{j2\pi f \tau_l}, \quad (2)$$

where $G(f)$ is the Fourier transform of $g(t)$. With proper guard subcarriers in OFDM systems, $G(f)$ can be regarded as a constant. Without loss of generality, assume $G(f) = 1$ in the subsequent discussion.

Under the uncorrelated scattering assumption [19], the complex amplitudes, α_l 's, can be modeled as independent complex Gaussian random variables. Hence, the correlation function of CFR is

$$r(\Delta f) \triangleq \mathbb{E}\{H^*(f + \Delta f)H(f)\} = \sum_l \sigma_l^2 e^{-j2\pi \Delta f \tau_l}, \quad (3)$$

where $\sigma_l^2 = \mathbb{E}(\alpha_l^2)$ is the power of the l -th path.

For an OFDM system with symbol duration, T , the subcarrier spacing will be $1/T$ according to the orthogonal condition. Given that there are K subcarriers, the CFR at the k -th subcarrier can be expressed as

$$H[k] \triangleq H(k/T). \quad (4)$$

Correspondingly, the correlation function for different subcarriers is described by

$$r[k] \triangleq r(k/T). \quad (5)$$

Define $\mathbf{H} \in \mathcal{C}^{K \times 1}$ to be the CFR vector, as

$$\mathbf{H} = \begin{pmatrix} H[0] \\ \vdots \\ H[K-1] \end{pmatrix}. \quad (6)$$

The correlation matrix of CFR vector, $\mathbf{H}$, is represented by

$$\mathbf{R}_\mathbf{H} \triangleq \mathbb{E}(\mathbf{H}\mathbf{H}^H) = \begin{pmatrix} r[0] & \cdots & r[1-K] \\ \vdots & \ddots & \vdots \\ r[K-1] & \cdots & r[0] \end{pmatrix}. \quad (7)$$

In a typical Rayleigh fading channel, the CFR vector, $\mathbf{H}$, is a complex Gaussian distributed random vector. Hence, the

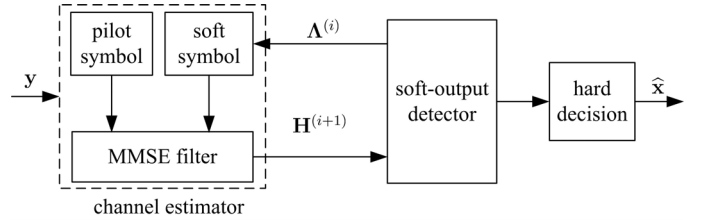


Fig. 1. A general structure for OFDM receiver with the ICE. The superscript, $(\cdot)^{(i)}$, denotes the i -th iteration, and $\Lambda^{(i)}$ is a $K \times 1$ vector, denoting the LLRs of K symbols.

probability density function (pdf) can be expressed as ¹ [19]

$$p(\mathbf{H}) = \frac{1}{\pi^K \det(\mathbf{R}_\mathbf{H})} e^{-\mathbf{H}^H \mathbf{R}_\mathbf{H}^{-1} \mathbf{H}}, \quad (8)$$

where $\det(\cdot)$ is the determinant of a matrix.

B. OFDM Receiver with Soft-Output Detector

In an OFDM system with K subcarriers, the received signal, after removing the *cyclic prefix* (CP) and *inverse fast Fourier transform* (IFFT), can be represented as

$$y[k] = H[k]x[k] + w[k], \quad (9)$$

where $y[k]$ denotes the received signal at the k -th subcarrier, $x[k]$ is the transmitted symbol at the k -th subcarrier, and $w[k]$ is the additive white Gaussian noise with zero mean and $\mathbb{E}(|w_k|^2) = \sigma_w^2$.

The transmitted symbol, $x[k]$, is either a pilot symbol that is known to the receiver or an unknown data symbol. The pilot symbols are used to estimate channel for the detection of the unknown data symbols. In this paper, we assume the transmitted symbols have constant modulus to facilitate our discussion. However, the proposed approach can be used to the cases of non-constant modulus symbols.

A general structure of OFDM receiver with the ICE is shown in Fig. 1. To perform channel estimation iteratively, soft information about unknown data symbols are fed back by a soft-output detector. Different from a hard detector, the soft-output detector yields not only the decision results but also the reliability of the decisions. Without loss of generality, we assume that four *quadrature amplitude modulation* (4-QAM) is used².

Denote $b[k, l]$, for $l = 0$ and 1 , to be the l -th bit of the k -th QAM symbol. Then, the modulated symbol can be expressed as

$$x[k] = \frac{\sigma_x}{\sqrt{2}} \left[(-1)^{b[k,0]} + j(-1)^{b[k,1]} \right], \quad (10)$$

and the *log-likelihood ratio* (LLR) for $b[k, l]$, yielded by the soft-output detector, can be represented as [8]

$$\Lambda[k, l] = \ln \frac{\mathbb{P}\{b[k, l] = 1 | \mathbf{y}, \mathbf{H}\}}{\mathbb{P}\{b[k, l] = 0 | \mathbf{y}, \mathbf{H}\}}, \quad (11)$$

¹Although $\mathbf{R}_\mathbf{H}$ is rank deficient and thus not invertible, we can only focus on the linear independent components of $\mathbf{H}$, and treat the other components as their deterministic linear combinations. This allows us to work with a compact notation in (8). Similar discussion about the rank deficient case can be found in Appendix A of [19].

²Following similar procedure, the discussion here can also be used for other modulation scheme, e.g. 16-QAM.

where $\mathbf{y} \in \mathcal{C}^{K \times 1}$ is the received signal vector and can be expressed as

$$\mathbf{y} = \begin{pmatrix} y[0] \\ \vdots \\ y[K-1] \end{pmatrix}. \quad (12)$$

With the LLR in (11), the conditional probabilities for $b[k, l] = 1$ and $b[k, l] = 0$, with $l = 0$ and 1 , can be expressed, respectively, as

$$P\{b[k, l] = 1 | \mathbf{y}, \mathbf{H}\} = \frac{e^{\Lambda[k, l]}}{1 + e^{\Lambda[k, l]}}, \quad (13a)$$

$$P\{b[k, l] = 0 | \mathbf{y}, \mathbf{H}\} = \frac{1}{1 + e^{\Lambda[k, l]}}. \quad (13b)$$

Therefore, the probability of $x[k] = c_{mn}$ is represented by

$$P\{x[k] = c_{mn} | \mathbf{y}, \mathbf{H}\} = P\{b[k, 0] = m | \mathbf{y}, \mathbf{H}\} P\{b[k, 1] = n | \mathbf{y}, \mathbf{H}\}, \quad (14)$$

where c_{mn} denotes the constellation point corresponding to $b[k, 0] = m, b[k, 1] = n$.

III. OPTIMUM ITERATIVE CHANNEL ESTIMATION

In this section, we will develop the optimum MAP estimation and discuss its implementation.

A. MAP Estimation with FPI

Given the received symbol vector, $\mathbf{y}$, the optimum estimation of CFR in terms of the MAP criterion can be expressed as

$$\begin{aligned} \hat{\mathbf{H}}_{\text{MAP}} &= \arg \max_{\{\mathbf{H}\}} \{\ln p(\mathbf{H} | \mathbf{y})\} \\ &= \arg \max_{\{\mathbf{H}\}} \{\ln p(\mathbf{H}) + \ln p(\mathbf{y} | \mathbf{H}) - \ln p(\mathbf{y})\}. \end{aligned} \quad (15)$$

Therefore, the MAP estimation of CFR satisfies

$$\frac{\partial \ln p(\mathbf{H})}{\partial \mathbf{H}^*} + \frac{\partial \ln p(\mathbf{y} | \mathbf{H})}{\partial \mathbf{H}^*} = \mathbf{0}, \quad (16)$$

where $(\cdot)^*$ denotes the conjugate. In (16), the first term corresponds to a priori knowledge about channel distribution, and the second term corresponds to the MLE estimation.

Using the pdf in (8), it is easy to obtain the first term in (16) as

$$\frac{\partial \ln p(\mathbf{H})}{\partial \mathbf{H}^*} = -\mathbf{R}_{\mathbf{H}}^{-1} \mathbf{H}. \quad (17)$$

If the transmitted symbols are known, such as pilot symbols, it is also easy to derive an analytical solution for the second term in (16). However, since there are unknown data symbols, the conditional pdf, $p(\mathbf{y} | \mathbf{H})$, is unknown, and thus it is difficult to find a solution directly. One straightforward way is to express the conditional pdf, $p(\mathbf{y} | \mathbf{H})$, as the sum of $p(\mathbf{y} | \mathbf{H}, \mathbf{x})$ using the complete-probability theorem. However, this approach makes the calculation complex and difficult for further manipulation.

To simplify it, we first rewrite the second term in (16) as

$$\begin{aligned} \frac{\partial \ln p(\mathbf{y} | \mathbf{H})}{\partial \mathbf{H}^*} &= \frac{1}{p(\mathbf{y} | \mathbf{H})} \cdot \frac{\partial p(\mathbf{y} | \mathbf{H})}{\partial \mathbf{H}^*} \\ &= \frac{1}{p(\mathbf{y} | \mathbf{H})} \sum_{q=1}^Q \frac{\partial p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q) P\{\mathbf{x}_q\}}{\partial \mathbf{H}^*}, \end{aligned} \quad (18)$$

where the second equation follows from the complete-probability theorem. In (18), $\mathbf{x}_q$ denotes the q -th possible symbol vector. For the QPSK modulation, we have $Q = 4^K$ possible symbol vectors if all the symbols are unknown data symbols. Since

$$\frac{\partial \ln p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q)}{\partial \mathbf{H}^*} = \frac{1}{p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q)} \cdot \frac{\partial p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q)}{\partial \mathbf{H}^*}, \quad (19)$$

then (18) can be written as

$$\begin{aligned} \frac{\partial \ln p(\mathbf{y} | \mathbf{H})}{\partial \mathbf{H}^*} &= \sum_{q=1}^Q \frac{p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q) P\{\mathbf{x}_q\}}{p(\mathbf{y} | \mathbf{H})} \cdot \frac{\partial \ln p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q)}{\partial \mathbf{H}^*} \\ &= \sum_{q=1}^Q P\{\mathbf{x}_q | \mathbf{y}, \mathbf{H}\} \cdot \frac{\partial \ln p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q)}{\partial \mathbf{H}^*}. \end{aligned} \quad (20)$$

Note that $P\{\mathbf{x}_q | \mathbf{y}, \mathbf{H}\}$ is the a posteriori probability of the q -th possible symbol vector, $\mathbf{x}_q$, and it depends on the detector we use. Given the q -th transmitted symbol vector, $\mathbf{x}_q$, and the CFR, $\mathbf{H}$, it is easy to obtain

$$p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q) = \frac{1}{(\pi \sigma_w^2)^K} e^{-\frac{1}{\sigma_w^2} \|\mathbf{y} - \mathbf{X}_q \mathbf{H}\|_2^2}, \quad (21)$$

where $\|\cdot\|_2$ denotes the 2-norm of a vector, and $\mathbf{X}_q$ is a $K \times K$ diagonal matrix with the k -th diagonal element from the k -th element of vector $\mathbf{x}_q$. Assuming constant modulus modulation is used, e.g. QPSK, we have

$$\frac{\partial \ln p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q)}{\partial \mathbf{H}^*} = \gamma (\mathbf{X}_q^{-1} \mathbf{y} - \mathbf{H}), \quad (22)$$

where $\gamma = \sigma_x^2 / \sigma_w^2$ is the SNR. Substituting (22) into (20), we finally obtain

$$\frac{\partial \ln p(\mathbf{y} | \mathbf{H})}{\partial \mathbf{H}^*} = \gamma (\bar{\mathbf{X}}^{-1}(\mathbf{H}) \mathbf{y} - \mathbf{H}) \quad (23)$$

where $\bar{\mathbf{X}}(\mathbf{H})$ is the soft symbol matrix given by

$$\bar{\mathbf{X}}(\mathbf{H}) = \left(\sum_{q=1}^Q P\{\mathbf{x}_q | \mathbf{y}, \mathbf{H}\} \mathbf{X}_q^{-1} \right)^{-1}. \quad (24)$$

Equation (24) implies that the soft symbol matrix is a function of the CFR.

Combining (17) and (22), the MAP estimation of CFR, $\mathbf{H}$, satisfies

$$\mathbf{H} = \mathbf{R}_{\mathbf{H}} \left(\mathbf{R}_{\mathbf{H}} + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \bar{\mathbf{X}}^{-1}(\mathbf{H}) \mathbf{y}. \quad (25)$$

It is impossible to find a closed-form solution for $\mathbf{H}$ from (25). The FPI approach [15] can be used to solve this kind of equation. With an initial estimation, $\mathbf{H}^{(0)}$, FPI solves (25)

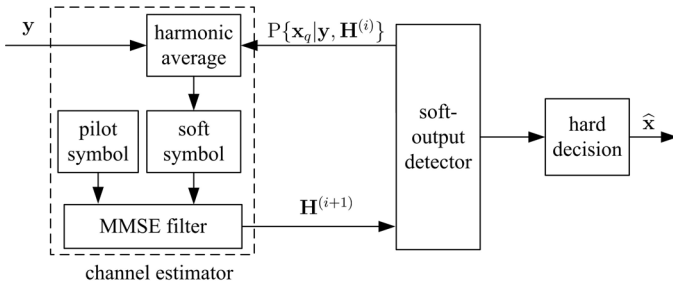


Fig. 2. The structural representation of the proposed MAP ICE.

iteratively using the following relation

$$\mathbf{H}^{(i+1)} = \mathbf{R}_H \left(\mathbf{R}_H + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \bar{\mathbf{X}}^{-1} (\mathbf{H}^{(i)}) \mathbf{y}, \quad (26)$$

for $i = 0, 1, 2, \dots$, where $\mathbf{H}^{(i)}$ denotes channel estimation after the i -th iteration. Usually, the initial estimation, $\mathbf{H}^{(0)}$, can be obtained using traditional pilot-aided channel estimation.

B. Harmonic Average

Note that the average in (24) is carried out over all possible symbol vectors. This is different from the traditional ICE, where the average is performed over possible constellation points for each symbol. Actually, the vector-wise average in (24) can be also simplified into a symbol-wise average by noting that the same symbol may occur in different symbol vectors. We consider the QPSK modulation as an example. The analysis can also be used for non-constant modulus modulation such as 16-QAM.

From (24), the k -th element of the soft symbol matrix, $\mathbf{X}(\mathbf{H}^{(i)})$, can be expressed as

$$\begin{aligned} \bar{x}[k] &= \left(\sum_{q=1}^Q P \left\{ \mathbf{x}_q | \mathbf{y}, \mathbf{H}^{(i)} \right\} x_q^{-1}[k] \right)^{-1} \\ &= \left(\sum_{q=1}^{Q/4} \sum_{m=0}^1 \sum_{n=0}^1 P \left\{ x[k] = c_{mn}, \mathbf{x}_{q,[k]} | \mathbf{y}, \mathbf{H}^{(i)} \right\} c_{mn}^{-1} \right)^{-1} \end{aligned} \quad (27)$$

where $\mathbf{x}_{q,[k]}$ denotes the q -th possible vector excluding the k -th element. Using the complete-probability theorem [20], we have

$$\sum_{q=1}^{Q/4} P \{ x[k] = c_{mn}, \mathbf{x}_{q,[k]} | \mathbf{y}, \mathbf{H}^{(i)} \} = P \{ x[k] = c_{mn} | \mathbf{y}, \mathbf{H}^{(i)} \}. \quad (28)$$

Then, (27) can be simplified as

$$\bar{x}[k] = \left(\sum_{m=0}^1 \sum_{n=0}^1 P \{ x[k] = c_{mn} | \mathbf{y}, \mathbf{H}^{(i)} \} c_{mn}^{-1} \right)^{-1} \quad (29)$$

where the probability $P \{ x[k] = c_{mn} | \mathbf{y}, \mathbf{H}^{(i)} \}$ is given in (14). As we can see, the proposed algorithm suggests a symbol-wise harmonic average for the data symbol, which is different from the arithmetic average widely used in the existing work.

A structural representation of the proposed MAP ICE is shown in Fig. 2. The MAP ICE is exactly the same as the OFDM receiver in Fig. 1 except that a harmonic average

is used for generating soft symbols. The soft symbols are regenerated using (29), and then sent to a *minimum mean-square-error* (MMSE) filter together with the pilot symbols, i.e. (26), to yield the channel estimation for the next iteration. For arithmetic average based approach [8], the soft symbol is regenerated as

$$\bar{x}[k] = \sum_{i=1}^M P \{ x[k] = c_i | \mathbf{y}, \mathbf{H} \} c_i. \quad (30)$$

When the SNR is small, there is little a posteriori information can be extracted, and thus we may have $P \{ x[k] = c_i | \mathbf{y}, \mathbf{H} \} \approx 1/M$ and in turn $\bar{x}[k] \approx 0$. Apparently, a zero-approaching soft symbol will degrade the channel estimation performance through $y[k]/\bar{x}[k]$. On the other side, there is no such problem if we use harmonic average. Hence, the performance using harmonic average can be improved especially at small SNRs as demonstrated by our simulation in Section VI.

In the above, we have presented the MAP ICE for QPSK modulation. For non-constant modulus modulation such as 16-QAM, we can use the approximation $\mathbf{X}\mathbf{X}^H \approx E(\mathbf{X}\mathbf{X}^H) = \sigma_x^2 \mathbf{I}$ where $E(|x[k]|^2) = \sigma_x^2$ is now the average transmit power. Then, the proposed approach above can also be used for non-constant modulus modulation.

C. Complexity

Since the constellation is known to the receiver, the inverse of constellation point, c_{mn}^{-1} , in (29) can be thus obtained in advance. We need $4K$ complex multiplication to yield the soft symbols during each iteration. Note that the complexity for harmonic average in (29) is the same as the arithmetic average in [8]. Taking into account the MMSE filter in (26), we need $K^2 + 4K$ complex multiplications in total during each iteration.

IV. CONVERGENCE ANALYSIS

In this section, we first discuss the relationship between the MAP ICE and EM estimation, and then investigate the convergence behavior.

A. Relation with EM Estimation

In the framework of EM estimation, the complete data, $\mathbf{z}$, contains the received symbol vector, $\mathbf{y}$, as the observation, and unknown data symbol vector, $\mathbf{x}$, as the missing data. CFR vector, $\mathbf{H}$, is unknown. The purpose is to estimate CFR in the presence of unknown data symbols.

Following the definition of EM estimation, the E-step can be represented as

$$\tilde{Q}(\mathbf{H} | \mathbf{H}^{(i)}) = \sum_{q=1}^Q P \{ \mathbf{x}_q | \mathbf{y}, \mathbf{H}^{(i)} \} \ln p(\mathbf{y}, \mathbf{x}_q | \mathbf{H}). \quad (31)$$

Using Bayesian theorem and noting that $\mathbf{x}_q$ and $\mathbf{H}$ are independent, we can obtain

$$\begin{aligned} \ln p(\mathbf{y}, \mathbf{x}_q | \mathbf{H}) &= \ln p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q) + \ln P \{ \mathbf{x}_q | \mathbf{H} \} \\ &= \ln p(\mathbf{y} | \mathbf{H}, \mathbf{x}_q) + \ln P \{ \mathbf{x}_q \}. \end{aligned} \quad (32)$$

Since the second term in (32), i.e. $\ln P\{\mathbf{x}_q\}$, is not used in the M-step, (31) is equivalent to

$$Q(\mathbf{H}|\mathbf{H}^{(i)}) = \sum_{q=1}^Q P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} \ln p(\mathbf{y}|\mathbf{H}, \mathbf{x}_q). \quad (33)$$

Then, given a priori distribution, $p(\mathbf{H})$, the M-step is

$$\begin{aligned} \mathbf{H}^{(i+1)} &= \arg \max_{\{\mathbf{H}\}} \left\{ \tilde{Q}(\mathbf{H}|\mathbf{H}^{(i)}) + p(\mathbf{H}) \right\} \\ &= \arg \max_{\{\mathbf{H}\}} \left\{ Q(\mathbf{H}|\mathbf{H}^{(i)}) + p(\mathbf{H}) \right\} \end{aligned} \quad (34)$$

which can be easily solved by setting the derivative to be zero. Direct calculation yields the same result as (26).

The analysis above shows that the proposed MAP ICE is equivalent to the EM estimation. Actually, the EM estimation has been widely adopted in OFDM receiver designs [21]–[23]. By carefully choosing the data set, the EM estimation can be used for channel estimation [21], [23] or signal detection [22]. The proposed MAP ICE takes into account the soft information fed back by the soft-output detector in the E-step. As a result, the ICE can be conducted with the a priori information of unknown data symbols, and thus the estimation performance can be further improved. This observation reveals the connection between the EM estimation and the widely adopted ICE in the turbo receiver [8], [9], where soft information is fed back to facilitate channel estimation. Instead of arithmetic average, the proposed MAP ICE shows that the best way to use soft information is the harmonic average.

B. Convergence of the MAP ICE

Since the MAP ICE can be explained in the framework of the EM estimation, the convergence of the proposed MAP ICE is guaranteed since the EM algorithm always converges with a proper initial value [16].

In view of (15), we define a cost function, $l(\mathbf{H})$, as

$$l(\mathbf{H}) = \ln p(\mathbf{H}) + \ln p(\mathbf{y}|\mathbf{H}). \quad (35)$$

As we have proved in Appendix A,

$$\begin{aligned} l(\mathbf{H}) - l(\mathbf{H}^{(i)}) &= \left[\tilde{Q}(\mathbf{H}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}) \right] - \\ &\quad \left[\tilde{Q}(\mathbf{H}^{(i)}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}^{(i)}) \right]. \end{aligned} \quad (36)$$

Since (34) is a maximization operation, the estimation for next iteration, $\mathbf{H}^{(i+1)}$, guarantees that

$$\tilde{Q}(\mathbf{H}^{(i+1)}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}^{(i+1)}) \geq \tilde{Q}(\mathbf{H}^{(i)}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}^{(i)}). \quad (37)$$

As a result, we have

$$l(\mathbf{H}^{(i+1)}) \geq l(\mathbf{H}^{(i)}), \quad (38)$$

which indicates that the MAP ICE approach converges monotonically.

C. Convergence Rate

The iterative procedure in (26) implicitly defines a mapping $\mathbf{H} \rightarrow \mathbf{g}(\mathbf{H})$ such that

$$\mathbf{H}^{(i+1)} = \mathbf{g}(\mathbf{H}^{(i)}). \quad (39)$$

Assuming that $\mathbf{H}^{(i)}$ converges to the true CFR and $\mathbf{g}(\mathbf{H})$ is differentiable at $\mathbf{H}$, which is true in our context, a Taylor expansion of (39) yields

$$\mathbf{H}^{(i+1)} - \mathbf{H} = \mathbf{J}_{\mathbf{H}} \cdot (\mathbf{H}^{(i)} - \mathbf{H}) + o(\|\mathbf{H}^{(i)} - \mathbf{H}\|_2^2), \quad (40)$$

where $o(\cdot)$ denotes the higher order infinitesimal, and $\mathbf{J}_{\mathbf{H}}$ is a $K \times K$ Jacobian matrix, defined by

$$\mathbf{J}_{\mathbf{H}} \triangleq \frac{\partial \mathbf{g}(\mathbf{H})}{\partial \mathbf{H}^T} = \mathbf{R}_{\mathbf{H}} \left(\mathbf{R}_{\mathbf{H}} + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \frac{\partial \bar{\mathbf{X}}^{-1}(\mathbf{H})\mathbf{y}}{\partial \mathbf{H}^T}. \quad (41)$$

In view of (40), the convergence rate is determined by the spectrum radius, $\rho(\mathbf{J}_{\mathbf{H}})$ [23].

If there is error-correction coding over the transmitted OFDM symbols, it is difficult to calculate the derivative in (41). However, we believe that redundancy in coding will help to improve channel estimation. For an uncoded transmission, the LLRs for the k -th symbol only depends on the received signal at the k -th subcarrier and the CFR at the k -th subcarrier. Then, the LLR in (11) can be simplified as [8]

$$\begin{aligned} \Lambda[k, l] &= \ln \frac{P\{b[k, l] = 1 | y[k], H[k]\}}{P\{b[k, l] = 0 | y[k], H[k]\}} \\ &= \begin{cases} \ln \frac{p_{10} + p_{11}}{p_{00} + p_{01}}, & \text{if } l = 0 \\ \ln \frac{p_{01} + p_{11}}{p_{00} + p_{10}}, & \text{if } l = 1 \end{cases} \end{aligned} \quad (42)$$

where p_{mn} is the conditional pdf given by

$$\begin{aligned} p_{mn} &= p(y[k] | H[k], x[k] = c_{mn}) \\ &= \frac{1}{\pi \sigma_w^2} e^{-\frac{1}{\sigma_w^2} |y[k] - H[k]c_{mn}|^2}. \end{aligned} \quad (43)$$

In (42), we have ignored the a priori information term since the channel estimator provides no a priori information to the detector.

In (42), the LLRs for the k -th symbol are independent of the CFRs at other subcarriers, so is the k -th soft symbol, $\bar{x}^{-1}[k]$. Hence, the derivative term in (41) becomes a diagonal matrix,

$$\frac{\partial \bar{\mathbf{X}}^{-1}(\mathbf{H})\mathbf{y}}{\partial \mathbf{H}^T} = \begin{pmatrix} \frac{\partial \bar{x}^{-1}[0]y[0]}{\partial H[0]} & & \\ & \ddots & \\ & & \frac{\partial \bar{x}^{-1}[K-1]y[K-1]}{\partial H[K-1]} \end{pmatrix}. \quad (44)$$

Using (29) and (42), after some manipulation, the inverse of the k -th soft symbol becomes

$$\bar{x}^{-1}[k] = \frac{1}{\sqrt{2}\sigma_x} \left[\left(\frac{1 - e^{\Lambda[k,0]}}{1 + e^{\Lambda[k,0]}} \right) + j \left(\frac{e^{\Lambda[k,1]} - 1}{e^{\Lambda[k,1]} + 1} \right) \right]. \quad (45)$$

The derivative of $\bar{x}^{-1}[k]$ with respect to $H[k]$ can be conducted for the real part and imaginary part, separately. Taking the derivative of the real part, $\text{Re}\{\bar{x}^{-1}[k]\}$, for example, the modulus of the real part for the k -th term of (44) can be expressed as

$$\left| \frac{\partial \text{Re}\{\bar{x}^{-1}[k]\}y[k]}{\partial H[k]} \right| = \left| \frac{\sqrt{2}y[k]}{\sigma_x (e^{\Lambda[k,0]} + e^{-\Lambda[k,0]} + 2)} \cdot \frac{\partial \Lambda[k,0]}{\partial H[k]} \right|. \quad (46)$$

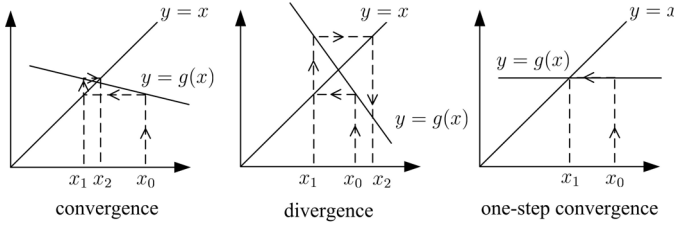


Fig. 3. Graphical explanation of the one step property.

As we have shown in Appendix B,

$$\left| \frac{\partial \Lambda[k, 0]}{\partial H[k]} \right| \leq \left| \frac{2\sigma_x y^*[k]}{\sigma_w^2} \right|. \quad (47)$$

Furthermore, from Appendix C, when the SNR is large enough, we have

$$|\Lambda[k, 0]| \geq 2|H[k]|^2\gamma. \quad (48)$$

Using the relation in (47) and (48), we can derive a high SNR approximation of (46) as

$$\left| \frac{\partial \text{Re}\{\bar{x}^{-1}[k]\} y[k]}{\partial H[k]} \right| \leq \frac{2\sqrt{2}|H[k]|^2\gamma}{e^{2|H[k]|^2\gamma} + e^{-2|H[k]|^2\gamma} + 2} \approx 0. \quad (49)$$

The second approximation is obtained by noting that the linear term of the SNR in the numerator is much smaller than the exponential term of the SNR in the denominator as the increase of the SNR. Similar conclusion can be drawn for the imaginary part, $\text{Im}\{\bar{x}^{-1}[k]\}$.

From (49), the Jacobian matrix, $\mathbf{J}_H$, approaches to zero when the SNR is high, and so is the spectrum radius, $\rho(\mathbf{J}_H)$. This means that the proposed approach has a superlinear convergence rate [23]. Actually, we can observe that the proposed algorithm can converge within one step from our simulation result in Section VI.

Recalling that the proposed approach is based on FPI, the one step convergence property can be also observed in a graphical manner in Fig. 3. For a scalar FPI solution, e.g. $x = g(x)$, as shown in the figure, the iterative procedure converges if and only if the absolute value of derivative at the true value, $|g'(x)|$, is smaller than unity [15]. Otherwise, it will diverge. A special case is that the iterative procedure can converge within one step when $g'(x) = 0$. For a vector case, the derivative, $g'(x)$, corresponds to the determinant of Jacobian matrix, i.e. $\det(\mathbf{J}_H)$. We have shown in the above that the Jacobian matrix approaches to zero when the SNR is large enough, and so is its determinant. Hence, we can expect that the proposed approach converges within one step.

V. PERFORMANCE BOUND

Since we have a priori information about the channel statistics, the performance bound can be evaluated using Bayesian CRB [18]. Different from the classical Bayesian CRB, the observation equation is partially known since the data symbols are unknown.

From [18], the MSE of channel estimation can be bounded by the inverse of *Bayesian information matrix* (BIM), $\mathbf{J}_B$, that

is

$$\text{MSE} \geq -\text{Tr}(\mathbf{J}_B^{-1}). \quad (50)$$

The BIM can be expressed as

$$\mathbf{J}_B = \mathbf{J}_P + \mathbf{J}_F, \quad (51)$$

where $\mathbf{J}_P$ is a priori information matrix, defined as

$$\mathbf{J}_P = \mathbf{E}_H \left[\frac{\partial^2 \ln p(\mathbf{H})}{\partial \mathbf{H}^* \partial \mathbf{H}^T} \right], \quad (52)$$

and, $\mathbf{J}_F$ is the *Fisher information matrix* (FIM), defined as

$$\mathbf{J}_F = \mathbf{E}_H \left\{ \mathbf{E}_{\mathbf{y}|\mathbf{H}} \left[\frac{\partial^2 \ln p(\mathbf{y}|\mathbf{H})}{\partial \mathbf{H}^* \partial \mathbf{H}^T} \right] \right\}. \quad (53)$$

From (17),

$$\mathbf{J}_P = -\mathbf{R}_H^{-1}. \quad (54)$$

Using (23), we can further derive that

$$\frac{\partial^2 \ln p(\mathbf{y}|\mathbf{H})}{\partial \mathbf{H}^* \partial \mathbf{H}^T} = -\gamma \mathbf{I} + \gamma \frac{\partial \bar{\mathbf{X}}^{-1}(\mathbf{H}) \mathbf{y}}{\partial \mathbf{H}^T}. \quad (55)$$

In (55), the first term indicates exactly the performance bound of the ML estimation when the transmitted symbols are all known pilots. The second term shows the impact of unknown data symbols. As we can see, the impact caused by unknown data symbols depends on the derivative of detector performance at CFR. Due to unknown data symbols, the second term presents, and thus the performance bound increases.

Further manipulation on the second term is difficult. However, we can gain some insights if the SNR is high enough. In this case, the second term becomes a diagonal matrix. Similar to (49), we can obtain the real part for the k -th diagonal element as

$$\left[\gamma \frac{\partial \text{Re}\{\bar{\mathbf{X}}^{-1}(\mathbf{H})\} \mathbf{y}}{\partial \mathbf{H}^T} \right]_{(k,k)} \approx \frac{2\sqrt{2}|H[k]|^2\gamma^2}{e^{2|H[k]|^2\gamma} + e^{-2|H[k]|^2\gamma} + 2} \approx 0. \quad (56)$$

Different from (49), the numerator in (56) is a square term of the SNR. Even though, it is still much smaller than the exponential term of the SNR in the denominator. Similarly, the imaginary part is also zero. As a result, the second term in (55) approaches to zero for a high SNR. Substituting the relations (54) to (56) into (50), we can obtain the expression of the MSE for high SNR, that is

$$\text{MSE} \geq \sum_{k=1}^K \frac{\lambda_k}{1 + \gamma \lambda_k}, \quad (57)$$

where λ_k is the k -th eigenvalue of the CFR correlation matrix, $\mathbf{R}_H$.

When SNR is high, the channel estimation is accurate enough. In this case, further improvement on channel estimation performance does not necessarily improve the detector performance. The detector performance is now mainly determined by the additive noise rather than the channel estimation error. In other words, we can say that the performance of detector is insensitive to the channel estimation error when the SNR is high. Mathematically, this insensitivity is equivalent to

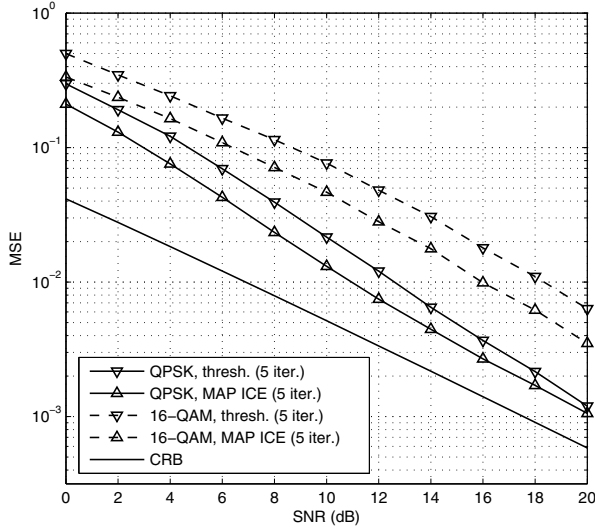


Fig. 4. MSE performance versus SNR with five iterations.

having a very small derivative at the CFR, as we have shown above.

VI. SIMULATION RESULTS

In this section, we present simulation results to demonstrate the performance of the proposed approach. For simulation, we consider an OFDM system with 128 subcarriers, with four guard subcarriers on each side. The subcarrier spacing is 15 KHz, corresponding to a symbol duration of about 0.067 msec. The CP is assumed long enough so that no inter-symbol interference in presence. QPSK and 16-QAM are both used in the simulation. Each OFDM symbol consists of 1/15 pilot uniformly distributed over the subcarriers and 14/15 data symbols. The 3GPP TU channel model is employed in the simulation [24], which has 20 paths and the maximum delay spread is 2.14 usec.

Fig. 4 shows the MSE performance of the proposed approach. The threshold control algorithm based on arithmetic average [10] is also included. Different from the typical arithmetic average based approach such as [8], threshold control approach can avoid zero-approaching soft symbols at small SNRs. As a result, the channel estimation performance can be improved at small SNRs. Although the threshold control based approach can avoid performance degradation at small SNR, it is just a sub-optimal approach. As a result, the proposed approach, based on MAP criterion, can outperform the threshold control based approach, especially at low-medium SNRs. As a benchmark, the CRB for high SNR, i.e. (57), is also included in the figure. Due to the presence of unknown data symbols, there is still a 2 – 3 dB performance gap for QPSK when the SNR is large.

Fig. 5 shows the MSE of the proposed approach with different iteration numbers. From the figure, the first iteration yields a significant performance improvement. More iterations can only yield small performance improvement. This coincides with our convergence analysis in Section IV. When the SNR is large enough, the proposed approach can converge within

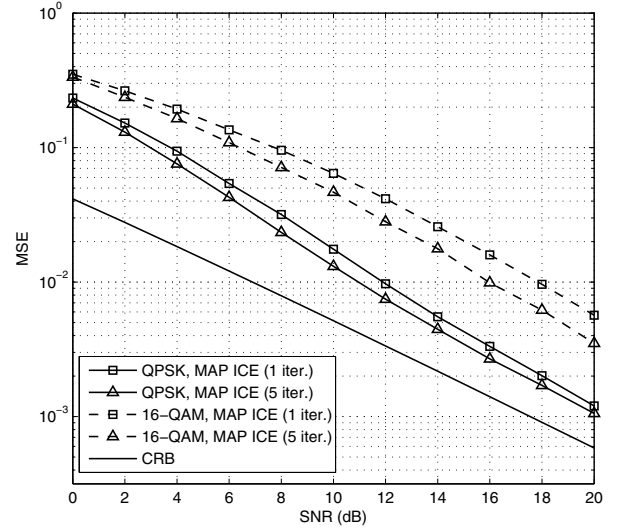


Fig. 5. MSE performance versus SNR with five iterations and one iteration.

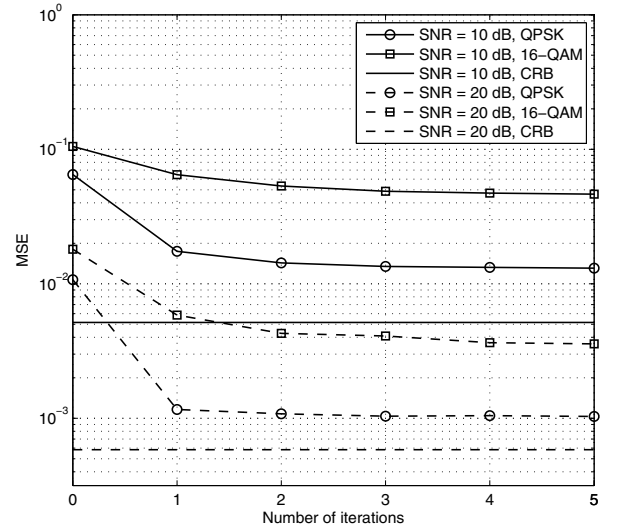


Fig. 6. MSE performance versus iteration numbers at 0 dB, 10 dB, and 20 dB, respectively.

one step, and therefore the performance is improved a lot for the first iteration. Since the algorithm has already converged, almost no improvement can be observed for the following iterations.

Fig. 6 further shows the one step convergence property of the proposed approach. When SNR = 20 dB, the first iteration yields a significant performance improvement while there is almost no improvement for further iterations. This is because the MAP ICE has converged after the first iteration for a high SNR. When the SNR is low, such as SNR = 10 dB, the second iteration still yields a little performance improvement since the approach does not converge yet.

Fig. 7 shows the symbol-error ratio (SER) versus SNR. The SER performances with perfect channel and estimated channel using threshold control based approach are also included for comparison. From the figure, the SER performance using

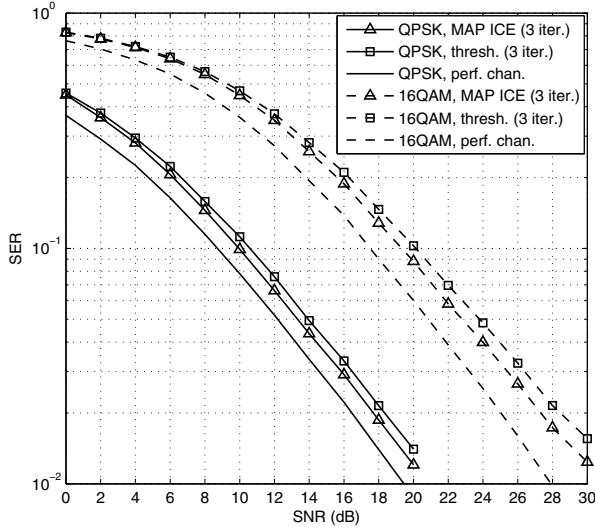


Fig. 7. SER performance versus SNR.

MAP ICE outperforms the one using threshold control based approach. For QPSK modulation, there is an about 0.7 dB performance between the MAP ICE approach and threshold control based approach. For 16-QAM, the performance gap is more than 1 dB. Higher order modulation is more sensitive to the accuracy of channel estimation. As a result, improved channel estimation accuracy can lead to more significant SER performance increase.

VII. CONCLUSIONS AND FURTHER RESEARCH

A MAP ICE approach is proposed in this paper to improve the channel estimation performance. The proposed approach is derived based on the MAP criterion, and uses a harmonic average over the constellation for regenerating the data symbol, and is different from the existing ICE approaches where arithmetic average is usually used. We have found that the proposed approach is closely related to the well-known EM estimation. This relation is utilized to study the convergence of the proposed approach. From the analysis and the simulation results, the proposed approach can significantly improve the performance within only one step.

Transmission over multiple OFDM symbols in time-varying channel will be also considered for further research. Moreover, the proposed approach will be extended to *multiple-input multiple-output* (MIMO) transmission.

APPENDIX A DERIVATION OF (36)

To derive (36), we first rewrite the cost function in (35) as

$$\begin{aligned}
 l(\mathbf{H}) &= \ln p(\mathbf{y}|\mathbf{H}) + \ln p(\mathbf{H}) \\
 &= \ln \sum_{q=1}^Q p(\mathbf{y}, \mathbf{x}_q|\mathbf{H}) + \ln p(\mathbf{H}) \\
 &= \ln \sum_{q=1}^Q \frac{p(\mathbf{y}, \mathbf{x}_q|\mathbf{H})}{P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\}} P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} + \ln p(\mathbf{H}).
 \end{aligned} \tag{A.1}$$

Using Jensen's inequality for the first term, we have

$$l(\mathbf{H}) \geq \sum_{q=1}^Q P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} \ln \frac{p(\mathbf{y}, \mathbf{x}_q|\mathbf{H})}{P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\}} + \ln p(\mathbf{H}). \tag{A.2}$$

With Bayesian theorem, we have

$$\ln \frac{p(\mathbf{y}, \mathbf{x}_q|\mathbf{H})}{P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\}} = \ln \frac{p(\mathbf{y}, \mathbf{x}_q|\mathbf{H})}{p(\mathbf{y}, \mathbf{x}_q|\mathbf{H}^{(i)})} p(\mathbf{y}|\mathbf{H}^{(i)}), \tag{A.3}$$

and thus (A.2) can be rewritten as

$$\begin{aligned}
 l(\mathbf{H}) &\geq \sum_{q=1}^Q P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} \ln \frac{p(\mathbf{y}, \mathbf{x}_q|\mathbf{H})}{p(\mathbf{y}, \mathbf{x}_q|\mathbf{H}^{(i)})} p(\mathbf{y}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}) \\
 &= \sum_{q=1}^Q P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} \ln p(\mathbf{y}, \mathbf{x}_q|\mathbf{H}) + \ln p(\mathbf{y}|\mathbf{H}^{(i)}) - \\
 &\quad \sum_{q=1}^Q P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} \ln p(\mathbf{y}, \mathbf{x}_q|\mathbf{H}^{(i)}) + \ln p(\mathbf{H})
 \end{aligned} \tag{A.4}$$

By appending a term, $\ln p(\mathbf{H}^{(i)}) - \ln p(\mathbf{H}^{(i)})$, to (A.4), it can be further rewritten as

$$\begin{aligned}
 l(\mathbf{H}) &\geq \left[\sum_{q=1}^Q P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} \ln p(\mathbf{y}, \mathbf{x}_q|\mathbf{H}) + \ln p(\mathbf{H}) \right] - \\
 &\quad \left[\sum_{q=1}^Q P\{\mathbf{x}_q|\mathbf{y}, \mathbf{H}^{(i)}\} \ln p(\mathbf{y}, \mathbf{x}_q|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}^{(i)}) \right] + \\
 &\quad \left[\ln p(\mathbf{y}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}^{(i)}) \right] \\
 &= \left[\tilde{Q}(\mathbf{H}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}) \right] - \\
 &\quad \left[\tilde{Q}(\mathbf{H}^{(i)}|\mathbf{H}^{(i)}) + \ln p(\mathbf{H}^{(i)}) \right] + l(\mathbf{H}^{(i)}).
 \end{aligned} \tag{A.5}$$

Move $l(\mathbf{H}^{(i)})$ in (A.5) to the left-hand side, then (36) is immediately obtained.

APPENDIX B DERIVATION OF (47)

From (42), we have

$$\left| \frac{\partial \Lambda[k, 0]}{\partial H[k]} \right| = \left| \frac{(p_{10} + p_{11}) \frac{\partial(p_{00} + p_{01})}{\partial H[k]} - (p_{00} + p_{01}) \frac{\partial(p_{10} + p_{11})}{\partial H[k]}}{(p_{10} + p_{11})(p_{00} + p_{01})} \right|. \tag{B.1}$$

By noting that

$$\frac{\partial p_{mn}}{\partial H[k]} = p_{mn} \left(\frac{y^*[k]c_{mn}}{\sigma_w^2} - H^*[k]\gamma \right), \tag{B.2}$$

(B.2) can be rewritten as

$$\begin{aligned}
 \left| \frac{\partial \Lambda[k, 0]}{\partial H[k]} \right| &= \left| \frac{y^*[k]}{\sigma_w^2} \right| \cdot \left| \frac{(p_{10} + p_{11})(p_{00}c_{00} + p_{01}c_{01}) - (p_{00} + p_{01})(p_{10}c_{10} + p_{11}c_{11})}{(p_{00} + p_{01})(p_{10} + p_{11})} \right| \\
 &= \left| \frac{y^*[k]}{\sigma_w^2} \right| \cdot \left| \frac{p_{00}c_{00} + p_{01}c_{01}}{p_{00} + p_{01}} - \frac{p_{10}c_{10} + p_{11}c_{11}}{p_{10} + p_{11}} \right|.
 \end{aligned} \tag{B.3}$$

Then, using the relation $|z_1 + z_2| \leq |z_1| + |z_2|$, where z_1, z_2 are any complex numbers, we can derive

$$\begin{aligned}
\left| \frac{\partial \Lambda[k, 0]}{\partial H[k]} \right| &\leq \left| \frac{y^*[k]}{\sigma_w^2} \right| \cdot \left(\left| \frac{p_{00}c_{00} + p_{01}c_{01}}{p_{00} + p_{01}} \right| + \left| \frac{p_{10}c_{10} + p_{11}c_{11}}{p_{10} + p_{11}} \right| \right) \\
&\leq \left| \frac{y^*[k]}{\sigma_w^2} \right| \cdot \left[\sigma_x \left(\frac{p_{00} + p_{01}}{p_{00} + p_{01}} \right) + \sigma_x \left(\frac{p_{10} + p_{11}}{p_{10} + p_{11}} \right) \right] \\
&= \left| \frac{2\sigma_x y^*[k]}{\sigma_w^2} \right| \quad (B.4)
\end{aligned}$$

APPENDIX C DERIVATION OF (48)

Using the approximation [25]

$$\ln(e^{\delta_1} + \dots + e^{\delta_n}) \approx \max_{i \in \{1 \dots n\}} \delta_i, \quad (C.1)$$

the LLR, $\Lambda[k, 0]$, can be approximated as

$$|\Lambda[k, 0]| = \left| \ln \frac{p_{10} + p_{11}}{p_{00} + p_{01}} \right| \approx \left| \ln \frac{\max\{p_{10}, p_{11}\}}{\max\{p_{00}, p_{01}\}} \right|. \quad (C.2)$$

Without loss of generality, we further assume $p_{10} = \max\{p_{10}, p_{11}\}$ and $p_{00} = \max\{p_{00}, p_{01}\}$. Then, (C.2) can be rewritten as

$$\begin{aligned}
|\Lambda[k, 0]| &\geq \left| \ln \frac{p_{10}}{p_{00}} \right| \\
&= \frac{1}{\sigma_w^2} |y^*[k]H[k]c_{10} + y[k]H^*[k]c_{10}^* - \\
&\quad y^*[k]H[k]c_{00} - y[k]H^*[k]c_{00}^*|. \quad (C.3)
\end{aligned}$$

Substituting (10) into (C.3), we have

$$|\Lambda[k, 0]| \geq \frac{2\sqrt{2}\sigma_x}{\sigma_w^2} |\operatorname{Re}\{y^*[k]H[k]\}|. \quad (C.4)$$

At high SNR, we ignore the noise term in the received signal, $y[k]$, then

$$|\operatorname{Re}\{y^*[k]H[k]\}| \approx |H[k]|^2 \cdot |\operatorname{Re}\{x[k]\}| = |H[k]|^2 \cdot \frac{\sigma_x}{\sqrt{2}}. \quad (C.5)$$

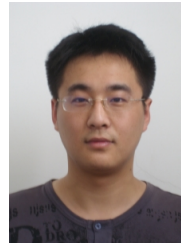
By substituting (C.5) into (C.4), we have

$$|\Lambda[k, 0]| \geq 2|H[k]|^2\gamma. \quad (C.6)$$

REFERENCES

- [1] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.
- [2] Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation, 3GPP Std., Rev. 36.211 (V9), Sept. 2008.
- [3] O. Edfors, M. Sandell, J. J. Beek, S. k. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, July 1998.
- [4] Y. G. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, July 2000.
- [5] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.
- [6] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. 1993 IEEE Int. Conf. Commun.*, pp. 1064–1070.
- [7] J. Hagenauer, "The turbo principle: tutorial introduction and state of the art," in *1997 Symp. Turbo Codes*.

- [8] F. Sanzi, S. Jeltin, and J. Speidel, "A comparative study of iterative channel estimators for mobile OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 849–859, Sept. 2003.
- [9] J. Bonnet and G. Auer, "Optimized iterative channel estimation for OFDM," in *Proc. 2006 IEEE Veh. Technol. Conf. – Fall*.
- [10] G. Auer and J. Bonnet, "Threshold controlled iterative channel estimation for coded OFDM," in *Proc. 2007 IEEE Veh. Technol. Conf. – Spring*, pp. 1737–1741.
- [11] M. C. Valenti and B. D. Woerner, "Refined channel estimation for coherent detection of turbo codes over flat fading," *Electron. Lett.*, vol. 34, no. 17, pp. 1648–1649, Aug. 1998.
- [12] H. Niu and J. A. Ritcey, "Iterative channel estimation and decoding of pilot symbol assisted LDPC coded QAM over flat fading channels," in *Proc. 2003 Asilomar Conf. Signals, Syst. Comput.*, vol. 2, pp. 2265–2269.
- [13] D. Yoon and J. Moon, "Low-complexity iterative channel estimation for turbo receivers," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1182–1187, May 2012.
- [14] C. T. Lam, D. D. Falconer, and F. D. Lemoine, "Iterative frequency domain channel estimation for DFT-precoded OFDM systems using in-band pilots," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 348–358, Feb. 2008.
- [15] T. Sauer, *Numerical Analysis (Chinese version)*. Posts and Telecom Press, 2010.
- [16] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood estimation from incomplete data via the em algorithm," *J. Royal Statistical Society (B)*, vol. 39, no. 1, pp. 1–38, 1977.
- [17] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Signal Process.*, vol. 36, no. 4, pp. 477–489, Apr. 1988.
- [18] H. V. Trees, "Bayesian Bounds." Available: http://www.ll.mit.edu/asap/asap_05/pdf/Presentations/01_vantrees.pdf
- [19] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [20] S. Goldberg, *Probability: An Introduction*. Dover, 1986.
- [21] X. Ma, H. Kobayashi, and S. C. Schwartz, "EM-based channel estimation algorithms for OFDM," *EURASIP J. Advances Signal Process.*, vol. 10, pp. 1460–1477, 2004.
- [22] B. Lu, X.-D. Wang, and Y. G. Li, "Iterative receivers for space-time block coded ofdm systems in dispersive fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 213–225, Apr. 2002.
- [23] Y. Xie and C. N. Georghiades, "Two EM-type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 106–115, Jan. 2003.
- [24] ETSI, "3GPP TR 25.943 version 9.0.0 Release 9." Available: <http://www.etsi.org/>
- [25] P. Roberston, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. 1995 IEEE Int. Conf. Commun.*, vol. 2, pp. 1009–1013.



Yinsheng Liu received the B.S. degree from North China Electric Power University, Baoding, China, and the M.S. degree from Beijing Jiaotong University, Beijing, China, in 2007 and 2009, respectively, all in communication and information system. He is currently pursuing his Ph.D degree at Beijing Jiaotong University, Beijing, China. His research interests include wireless communication system and digital signal processing.



Geoffrey Ye Li received his B.S.E. and M.S.E. degrees in 1983 and 1986, respectively, from the Department of Wireless Engineering, Nanjing Institute of Technology, Nanjing, China, and his Ph.D. degree in 1994 from the Department of Electrical Engineering, Auburn University, Alabama.

He was a Teaching Assistant and then a Lecturer with Southeast University, Nanjing, China, from 1986 to 1991, a Research and Teaching Assistant with Auburn University, Alabama, from 1991 to 1994, and a Post-Doctoral Research Associate with the University of Maryland at College Park, Maryland, from 1994 to 1996. He was with AT&T Labs - Research at Red Bank, New Jersey, as a Senior and then a Principal Technical Staff Member from 1996 to 2000. Since 2000, he has been with the School of Electrical and Computer Engineering at Georgia Institute of Technology as an Associate and then a Full Professor. He is also holding the Cheung Kong Scholar title at the University of Electronic Science and Technology of China since March 2006.

His general research interests include statistical signal processing and telecommunications, with emphasis on cross-layer optimization for spectral- and energy-efficient networks, cognitive radios, and practical techniques in LTE systems. In these areas, he has published over 300 referred journal and conference papers in addition to 20 granted patents. His publications have been cited over 15,000 times and he is listed as a highly cited researcher by Thomson Reuters. He once served or is currently serving as an editor, a member of editorial board, and a guest editor for over 10 technical journals. He organized and chaired many international conferences, including technical program vice-chair of IEEE ICC'03, technical program co-chair of IEEE SPAWC'11, and general chair of IEEE GlobalSIP'14. He has been awarded an IEEE Fellow for his contributions to signal processing for wireless communications since 2006, selected as a Distinguished Lecturer for 2009 - 2010 by IEEE Communications Society, and won 2010 Stephen O. Rice Prize Paper Award from IEEE Communications Society in the field of communications theory and 2013 James Evans Avant Garde Award from IEEE Vehicular Technology Society for advancing the state-of-art in OFDM-aided wireless communications.



Hongjie Hu received the B. Eng. degree from Northwestern Polytechnic University, China, and M. Eng. degree from the National University of Singapore, Singapore, both in electrical engineering. He has been with Huawei Technologies for 10 years, and works on air interface algorithms design and radio network performance analysis for various types of cellular base stations



Zhenhui Tan Professor and Ph.D. in Communications and Information Systems, was born on Feb. 18, 1944 in Yangzhou, Jiangsu Province. He graduated from Beijing Jiaotong University (BJTU) and majored in wireless communications in 1967. He received his master's degree in 1981 at BJTU and Ph.D. in 1987 at Southeast University, respectively. From 1998 to 2008, Professor Tan was the President of BJTU. From 1995 to 1998, Professor Tan was the Vice President of BJTU. From 1993 to 1995, Professor Tan was the Director of Department of Communications and Signal Control Engineering in BJTU. From 1990 to 1993, Tan studied as visiting scholar at Mons Technology College, Belgium and at University of Waterloo, Canada. Professor Tan is a member of the National Hi-Tech Project's Specialist Committee for three consecutive tenures, A member of the Evaluation Committee for Academic Degree Affairs of the State Council for two consecutive tenures, A member of Specialist Committee for Postdoctoral Degree Affairs, Director of electrical information science and engineering teaching guidance committee of Ministry of Education. Professor Tan was awarded by State Excellent Coming-Back From-Abroad Scholars, Experts of Outstanding Achievements from the State Council.

Professor Tan's research interest is mainly in communications and information systems, including digital mobile communications, spread spectrum communications and adaptive filtering algorithms, etc.